

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:38:27 ; Search time 6228 Seconds  
(without alignments)  
11467.291 Million cell updates/sec

Title: US-09-973-827-3  
Perfect score: 2454  
Sequence: 1 cggcgagggtgtgtttg.....tcgtattgtgtatcatc 2454

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 841850

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hgt: \*  
3: gb\_in: \*  
4: gb\_em: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pi: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_ov: \*  
22: em\_ph: \*  
23: em\_pat: \*  
24: em\_pi: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_un: \*  
28: em\_vi: \*  
29: em\_vt: \*  
30: em\_hgt\_hum: \*  
31: em\_hgt\_inv: \*  
32: em\_hgt\_other: \*  
33: em\_hgt\_mus: \*  
34: em\_hgt\_pln: \*  
35: em\_hgt\_rod: \*  
36: em\_hgt\_mam: \*  
37: em\_hgt\_vrt: \*  
38: em\_sy: \*  
39: em\_hgtg\_hum: \*  
40: em\_hgtg\_mus: \*  
41: em\_hgtg\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	27	1.1	27	6 AR090566	AR090566 Sequence
C 2	27	1.1	27	6 AR197601	AR197601 Sequence
C 3	26	1.1	26	6 AR090565	AR090565 Sequence
C 4	26	1.1	26	6 AR197600	AR197600 Sequence
C 5	25	1.0	49	1 ECOINSM	M11752 E.coli Inse
C 6	23.2	0.9	50	10 NMU41943	U41943 Mus musculus
C 7	23	0.9	50	6 AX161232	AX161232 Sequence
C 8	22.4	0.9	50	6 AX157048	AX157048 Sequence
C 9	22.2	0.9	43	6 AX183407	AX183407 Sequence
C 10	22	0.9	47	6 AX319484	AX319484 Sequence
C 11	22	0.9	48	6 AX319485	AX319485 Sequence
C 12	21.4	0.9	41	6 AX045523	AX045523 Sequence
C 13	21.2	0.9	45	6 AX045511	AX045511 Sequence
C 14	21.2	0.9	48	6 AX223683	AX223683 Sequence
C 15	21	0.9	43	6 AX484546	AX484546 Sequence
C 16	21	0.9	50	6 AR022597	AR022597 Sequence
C 17	21	0.9	50	6 AR037612	AR037612 Sequence
C 18	21	0.9	50	6 AR178070	AR178070 Sequence
C 19	21	0.9	50	6 AX160550	AX160550 Sequence
C 20	20.8	0.8	43	6 AX166164	AX166164 Sequence
C 21	20.8	0.8	43	6 E33229	E33229 Novel metha
C 22	20.8	0.8	45	6 AX320846	AX320846 Sequence
C 23	20.8	0.8	45	6 AX320847	AX320847 Sequence
C 24	20.8	0.8	48	6 AX428511	AX428511 Sequence
C 25	20.8	0.8	48	9 HS295110	295110 H.sapiens m
C 26	20.8	0.8	49	6 AX456410	AX456410 Sequence
C 27	20.6	0.8	36	6 E14547	E14547 DNA probe f
C 28	20.6	0.8	40	6 AR149458	AR149458 Sequence
C 29	20.6	0.8	40	6 E49430	E49430 Method for
C 30	20.6	0.8	43	6 AX484404	AX484404 Sequence
C 31	20.6	0.8	44	6 AR172650	AR172650 Sequence
C 32	20.6	0.8	45	9 HUMAIDSB1	M60135 Human/HIV 5
C 33	20.6	0.8	48	6 AX223789	AX223789 Sequence
C 34	20.6	0.8	48	6 AX426775	AX426775 Sequence
C 35	20.6	0.8	50	6 AR120175	AR120175 Sequence
C 36	20.6	0.8	50	6 AR126173	AR126173 Sequence
C 37	20.6	0.8	50	6 AR177989	AR177989 Sequence
C 38	20.6	0.8	50	6 AX061557	AX061557 Sequence
C 39	20.4	0.8	40	6 AR013538	AR013538 Sequence
C 40	20.4	0.8	40	6 AR109499	AR109499 Sequence
C 41	20.4	0.8	40	6 I55670	I55670 Sequence 73
C 42	20.4	0.8	40	6 I76447	I76447 Sequence 23
C 43	20.4	0.8	48	6 AR032406	AR032406 Sequence
C 44	20.4	0.8	48	6 AR209070	AR209070 Sequence
C 45	20.4	0.8	48	6 AX223778	AX223778 Sequence

ALIGNMENTS

RESULT 1  
AR090566/c  
LOCUS AR090566  
DEFINITION Sequence 686 from patent US 5994076.  
ACCESSION AR090566  
VERSION AR090566.1 GI:10017321  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 27)  
Chenchik, A., Johadze, G. and Bibilashvili, R.  
TITLE Methods of assaying differential expression  
JOURNAL Patent: US 5994076-A 686 30-NOV-1999;  
FEATURES Location/Qualifiers

27 bp DNA linear PAT 07-SEP-2000

```
source 1. .27
BASE COUNT 4 a 4 c 10 g 9 t
ORIGIN
Query Match 1.1% Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 827 ATACGAGATTCGCACACCCACTAG 853
Db 27 ATACGAGATTCGCACACCCACTAG 1

RESULT 2
AR197601/c
LOCUS AR197601 27 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 686 from patent US 6352829.
ACCESSION AR197601
VERSION AR197601.1 GI:20247450
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 6352829-A 636 05-MAR-2002;
FEATURES Location/Qualifiers
source 1. .26
BASE COUNT 4 a 4 c 10 g 9 t
ORIGIN
Query Match 1.1% Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 827 ATACGAGATTCGCACACCCACTAG 853
Db 27 ATACGAGATTCGCACACCCACTAG 1

RESULT 3
AR090565
LOCUS AR090565 26 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 685 from patent US 5994076.
ACCESSION AR090565
VERSION AR090565.1 GI:10017320
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 26)
AUTHORS Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 5994076-A 685 30-NOV-1999;
FEATURES Location/Qualifiers
source 1. .26
BASE COUNT 9 a 5 c 9 g 3 t
ORIGIN
Query Match 1.1% Score 26; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 7.7e+05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 507 GCACCAGAGTGCCAGGATTGAAGA 532
Db 1 GCACCAGAGTGCCAGGATTGAAGA 26

RESULT 4
AR197600
LOCUS AR197600 26 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 685 from patent US 6352829.
ACCESSION AR197600
VERSION AR197600.1 GI:20247449
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 26)
AUTHORS Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 6352829-A 685 05-MAR-2002;
FEATURES Location/Qualifiers
source 1. .26
BASE COUNT 9 a 5 c 9 g 3 t
ORIGIN
Query Match 1.1% Score 26; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 7.7e+05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 507 GCACCAGAGTGCCAGGATTGAAGA 532
Db 1 GCACCAGAGTGCCAGGATTGAAGA 26

RESULT 5
ECOINSW
LOCUS ECOINSW 49 bp DNA linear BCT 26-APR-1993
DEFINITION E.coli insertion site for transposon Tn1545.
ACCESSION M31752
VERSION M31752.1 GI:146490
KEYWORDS
SOURCE E.coli DNA.
ORGANISM Escherichia coli
REFERENCE 1 (bases 1 to 49)
AUTHORS Caillaud,F. and Courvalin,P.
TITLE Nucleotide sequence of the ends of the conjugative shuttle
transposon Tn1545
JOURNAL Mol. Gen. Genet. 209 (1), 110-115 (1987)
MEDLINE 88038347
PUBMED 2823065
FEATURES Location/Qualifiers
source 1. .49
BASE COUNT 12 a 10 c 0 g 27 t
ORIGIN
Query Match 1.0% Score 25; DB 1; Length 49;
Best Local Similarity 75.6%; P-ed. No. 1.1e+06;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 2300 CAATATTGGATTGTCATTCTTACAAACATTTTTCGTC 2340
Db 8 CAATTTCTCTTTTATTATTATAAAATCATTTTTCCTC 48

RESULT 6
MMU41943/c
LOCUS MMU41943 50 bp DNA linear ROD 05-JAN-1996
DEFINITION Mus musculus recombination between immunoglobulin heavy chain and
c-myc.
ACCESSION U41943
VERSION U41943.1 GI:1147577
KEYWORDS
SOURCE house mouse strain-BALB/can.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

ORGANISM Homo sapiens  
Euleria, Euteleostomi;  
Mammalia, Euterii;  
1 (bases 1 to 50)  
AUTHORS Shinkov, R. A. and Leach, M.  
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof  
JOURNAL patent: WO 0140521-A 376 07-JUN-2001

JOURNAL Patent: WO 0140541-A 376 07-JUN-2001;  
Curagen Corporation (US)  
FEATURES  
Source localism/Qualifiers  
1 50  
/organism="Homo sapiens"  
/db-xref="taxon:9606"  
misc-feature 25..26  
/note="Nucleotide deleted between bases 25 and 26"  
Accession number CG44921986

```

PRASE COUNT      15 a 3 c 6 g 26 t
ORIGIN
Query Match      0.9%  Score 22.4;  DB 6;  Length 50;
Best Local Similarity 65.7%;  Pred No. 3.7e+06;
Matches 32;  Conservative 0;  Mismatches 16;  Indels 0;  Gaps 0;
/note="2 of 2 allelic variants (375 is other entry)"

```

[illegible]

AX483407.1 GI:22317827

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

*Candida albicans*.  
*Candida albicans*.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Mitosporic Saccharomycetales; *Candida*.  
1. Roemer, T., Jlang, B., Boone, C., Bussey, H. and Ohlsen, K. L.  
Genes disrupting methodologies for drug target discovery  
Patent: WO 0203718-A 707 11-JUL-2002;  
http://www.patent.gov.uk/wipo/pubs/wo/wo0203718a.pdf

```

FEATURES
  source
    1. .43
      Location/Qualifiers
        /organism="Candida albicans"
        /db_xref="taxon:5476"
BASE COUNT      6 a      13 c      1 g      23 t
ORIGIN
Query Match      0.98; Score 22.2; DB 6; Length 43;

```

	Matches	30;	Conservative	0;	Mismatches	13;	Indels	0;	Gaps	0;
ay	152	TGAGATCGACGTCTACACAGAAGCTGGAAACCACCAACAATGACA								194
b	43	TGAGCTTCAACTCTGGAAGAACAGCAAGGAGGAAAAATGAAA								1

RESULT 10

LOCUS: AX319484  
DEFINITION: Sequence 13 from Patent WO0182962.  
ACCESSION: AX319484  
VERSION: AX319484.1 GI:17901270  
KEYWORDS: Human immunodeficiency virus.  
SOURCE: Human immunodeficiency virus.  
ORGANISM: HIV-1  
TAXID: 11068  
OTHER: PAT 14-DEC-2000 linear DNA 47 bp

```

/db_xref="taxon:32630"
/note="Synthetic"
19 a      4 c      5 g      13 t

Query Match      0.9%; Score 21.4; DB 6; Length 41;
Best Local Similarity 71.8%; Pred. No. 5.9e+06;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1855 ATTGTAATAAGCAAGTAAATATAAAGTACAAAGCAT 1893
||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
Ddb 1 ATGTTAGTACAAATAAGTATATGAATTTCAAATCAT 39

RESULT 13
LOCUS A19511 45 bp DNA linear PAT 10-JUN-1994
DEFINITION Oligonucleotide 61-83 Hepatitis A virus.
ACCESSION A19511
VERSION A19511.1 GI:583227
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 45)
AUTHORS Brown,A.L., Clarke,B.E. and Rowlands,D.J.
TITLE Chimeric hepatitis virus core antigen proteins
JOURNAL Patent: EP 0421635-A 37 10-APR-1991;
THE WELLCOME FOUNDATION LIMITED
FEATURES
Source 1..45
/organism="synthetic construct"
/db_xref="taxon:32630"
11 a 11 c 13 g 10 t

BASE COUNT 11 a 11 c 13 g 10 t
ORIGIN

Query Match      0.9%; Score 21.2; DB 6; Length 45;
Best Local Similarity 69.0%; Pred. No. 6.3e+06;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 604 CTAGCAGTGGACATATATGCCATTACCCAGGAGGAGCAA 645
||||| 11 ||||| 11 ||||| 11 ||||| 11 |||||
Ddb 2 CTAGCACTGAACAGAAATGTTCCGATCCTCAGTTGGAGCTA 43

RESULT 14
LOCUS AX223683 48 bp mRNA linear PAT 07-SEP-2001
DEFINITION Sequence 9125 from Patent WO0159103.
ACCESSION AX223683
VERSION AX223683.1 GI:15551407
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 48)
AUTHORS Blatt,L., Meswiggen,J. and Chowrira,B.M.
TITLE Method and reagent for the modulation and diagnosis of cd20 and
JOURNAL nogo gene expression
PATENT: WO 0159103-A 9125 16-AUG-2001;
RIBOZYNE PHARMACEUTICALS, INC. (US); Blatt, Lawrence (US);
Meswiggen, James (US); Chowrira, Bharat M. (US)
FEATURES
Source 1..48
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Nucleic Acid"
12 a 12 c 16 g 8 t

BASE COUNT 12 a 12 c 16 g 8 t
ORIGIN

Query Match      0.9%; Score 21.2; DB 6; Length 48;
Best Local Similarity 69.0%; Pred. No. 6.3e+06;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

SECRETARY

Oy 443 GCGAGGGAAATTC TTTC AAGGAGGCCTT CCTACAGGAAAT 484  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 7 GCGAGGAAACTCCCTTCAAGGACATCG TCGGGAGGAAAGT 48

RESULT 15					
AX484546/c					
LOCUS	AX484546	43 bp	DNA	linear	PAT 16-AUG-2002
DEFINITION	Sequence	1846 from Patent WO02053728.			
ACCESSION	AX484546				
VERSION	AX484546.1	GI:22318898			

**KEYWORDS**  
SOURCE  
ORGANISM  
Candida albicans.  
Candida albicans  
Eukaryota; Fungi; Ascomycota: Saccharomycotina: Saccharomycetes: Saccharomycetales; mitosporic Saccharomycetales; Candida.

**REFERENCE**  
1 Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L. Gene disruption methodologies for drug target discovery Patent: WO 02053728-A 1846 11-JUL-2002; Elutra Pharmaceuticals, Inc. (US)

```

FEATURES
source
Location/Qualifiers
1..43
/organism="Candida albicans"
/db_xref="taxon:5476"
BASE COUNT      12 a      6 c      4 g      21 t

```

Query Match 0.9%; Score 21; DB 6; Length 43;  
Best Local Similarity 73.0%; Pred. No. 7e+06;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY	1859	TAPATTAAGCRAAGTAAATAAAAGTACAAGCATAT	1895
Db	41	TAGITTAACGCATATTAAAAAAGAAATAAGCTTAT	5

Search completed: March 9, 2003, 20:52:22  
Job time : 6232 secs

OM nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:19:02 : Search time 510 Seconds  
(without alignments)  
10836.086 Million cell updates/sec

Title: US-09-973-827-3

Perfect score: 2454  
Sequence: 1 cgcggcggagggtggtgtg.....tctgtatttggatatcat 2454

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2166140

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_10:002:.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
C 1	27	1.1	27	24	ABK65598	Human gene specific
C 2	26.6	1.1	47	21	AA267081	Human map-related
C 3	26	1.1	26	24	ABK65597	Human gene specific
C 4	24	1.0	34	24	ABL40868	CREB cDNA amplifi
C 5	24	1.0	34	24	ABL40870	CREB cDNA amplifi
C 6	23.4	1.0	31	19	AAV21495	Plasmid PRSV-CREB3
C 7	23.4	1.0	40	21	AA266915	Human map-related
C 8	23	0.9	50	22	AA177619	Human silent SNP c
C 9	22.6	0.9	50	22	AA134381	Human SNP oligonuc

c 10	22.4	0.9	50	22	AA173435	Human silent SNP c
c 11	22.2	0.9	45	17	AA07082	Primer EXDO.1, amp
c 12	22	0.9	33	24	ABL40867	CREB cDNA amplifi
c 13	22	0.9	46	24	ABL01753	Human MSH2 (HMSH2)
c 14	22	0.9	47	24	AA171996	VCPI579 primer H1V
c 15	22	0.9	48	24	AA171997	VCPI579 primer H1V
c 16	21.6	0.9	47	21	AA267881	Human map-related
c 17	21.6	0.9	49	20	AA350326	Trailer oligonucle
c 18	21.4	0.9	40	13	AAQ25032	Oligonucleotide sp
c 19	21.4	0.9	41	21	AA087847	Bacillus thuringie
c 20	21.4	0.9	50	22	AA128587	Human SNP oligonuc
c 21	21.4	0.9	50	22	AA128695	Human SNP oligonuc
c 22	21.2	0.9	30	20	AA176677	CREB341 phosphor
c 23	21.2	0.9	45	12	AAQ11255	Encodes Hepatitis
c 24	21.2	0.9	48	23	AAK09125	Human SNP oligonuc
c 25	21.2	0.9	50	22	AAK129818	Human CD20 Ambery
c 26	21	0.9	32	19	AAV21501	Plasmid PRSV-CREB3
c 27	21	0.9	33	20	AA176678	CREB341 phosphor
c 28	21	0.9	47	21	AA269260	Human map-related
c 29	21	0.9	50	22	AA176937	Human silent SNP c
c 30	20.8	0.8	43	21	AA261260	Primer BAYqk-Gl fo
c 31	20.8	0.8	45	21	AA255674	HaEPV transfer vec
c 32	20.8	0.8	47	21	AA269043	Human map-related
c 33	20.8	0.8	48	24	A3K22203	Human ERG ambery
c 34	20.8	0.8	49	24	A3L99285	Human ERG ambery
c 35	20.8	0.8	49	24	A3L99285	Synthetic Renilla
c 36	20.8	0.8	50	22	AA134100	Human SNP oligonuc
c 37	20.6	0.8	36	19	AA134577	Human SNP oligonuc
c 38	20.6	0.8	40	22	AA269662	Nucleotide sequenc
c 39	20.6	0.8	44	21	AA290218	Human c-fos strept
c 40	20.6	0.8	44	24	ABL36027	Wild-type phoA pro
c 41	20.6	0.8	44	24	ABL36030	A. thaliana connec
c 42	20.6	0.8	48	23	ABK09231	A. thaliana connec
c 43	20.6	0.8	48	24	ABK22464	Human CD20 Ambery
c 44	20.6	0.8	48	24	ABK22464	Human ERG ambery
c 45	20.6	0.8	50	11	AAQ06172	Probe derived from
c 45	20.6	0.8	50	14	AAQ37565	RP10 derived BMP-6

ALIGNMENTS

RESULT 1  
ABK65598/c  
ID ABK65598 standard; DNA: 27 BP.

XX AC ABK65598;  
XX DT 02-JUL-2002 (first entry)  
XX DE Human gene specific PCR primer #686.  
XX KW Primer; ss; DNA microarray; differential expression analysis; human.  
XX OS Homo sapiens.  
XX PN US6352829-B1.  
XX PD 05-MAR-2002.  
XX PF 05-JAN-1999; 99US-0225928.  
XX PR 21-MAY-1997; 97US-0859998.  
XX PA (CLON-) CLONTECH LAB INC.  
XX PI Chenchik A, Johhadze G, Bibilashvilli R;  
XX DR WPT; 2002-314699/35.  
XX PT Producing sub-population of labeled nucleic acids, useful for analysing  
XX PT differences in RNA profiles between several different physiological  
XX PT sources, using set of distinct gene specific primers

WPI; 2000-013267/01.

Novel biallelic markers used to construct a high density disequilibrium map of the human genome

Claim 1; Page 535; 2745pp; English.

AAZ65654 to AAZ69578 represent human biallelic markers from the present invention, which contain a polymorphic base at position 24 of their nucleotide sequences. AAZ69579 to AAZ77440 represent amplification primers for the biallelic markers. The biallelic markers of the invention have a variety of uses: they can be used for high density mapping of the human genome, and in complex association studies and haplotyping studies which are useful in determining the genetic basis for disease states. Compositions and methods of the invention can also be useful for the identification of the targets for the development of pharmaceutical agents and diagnostic methods, as well as the characterisation of the differential efficacious responses to and side effects from pharmaceutical agents acting on a disease as well as other treatment.

N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and 3367, are not actually given a sequence in the Sequence Listing from the present invention.

Sequence 47 BP; 15 A; 4 C; 7 G; 21 T; 0 other;

Query Match 1.1%; Score 26.6; DB 21; Length 47;  
Best Local Similarity 78.0%; P-red. No. 2.1e+04;  
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1889 ACACATTTTGTAGTAGCTACTTAATTCCTTGTGATAAATGCCTGA 1929  
| | | | | | | | | | | | | | | | | | | | |  
DB 43 AACACATTTTGTAGTAGCTACTTAATTCCTTGTGATAAATGCCTGA 3

RESULT 3  
ABK66597  
ID ABK66597 standard; DNA; 26 BP.  
XX AC  
XX AC ABK66597;  
XX XX  
DT 02-JUL-2002 (first entry)  
XX DE  
XX Human gene specific PCR primer #685.  
KW Primer; ss; DNA microarray; differential expression analysis; human.  
XX Homo sapiens.  
OS US6352829-B1.  
PN PN  
XX XX  
PD 05-MAR-2002.  
XX XX  
PF 05-JAN-1999; 99US-0225928.  
XX XX  
PR 21-MAY-1997; 97US-0859998.  
XX XX  
PA (CLON-) CLONTECH LAB INC.  
XX XX  
PI Chenchik A, Jakhade G, Bibilashvili R;  
XX WPI; 2002-314699/35.  
XX XX  
PT Producing sub-population of labeled nucleic acids, useful for analysing  
PT differences in RNA profiles between several different physiological  
PT sources, using set of distinct gene specific primers  
XX XX  
PP Example 3; SEQ ID NO 685; 1lpp; English.  
XX XX  
XX The invention relates to producing a sub-population of labeled nucleic  
XX acids (NAs) comprising contacting a NA sample from a physiological  
CC source, with a pool of 50 distinct gene specific primers under suitable

CC  
CC  
CC  
CC  
The invention relates to a method of detecting the target gene of a specific transcription factor. The method involves (a) supplying cells

The invention relates to a method of detecting the target gene of a specific transcription factor. The method involves (a) supplying calls containing a vector expressing a fused protein of the transcription factor with an activator of this transcription factor and another vector containing a reporter gene; (b) separating cells; and (c) comparing the gene expression dose of the cells. The method is for identifying a target gene of transcription factor for isolation, which is used for application in cancer diagnosis and drug development for e.g. Li-Fraumeni syndrome and long-term memory loss. With expression of the reporter gene as indication, the identification and isolation can be efficiently carried out because its enlarged expression dose in a cell is induced by the transcription factor. Sequences ABL0867-874 represents primers used in the course of the invention.



XX SQ Sequence 34 BP; 10 A; 4 C; 10 G; 10 T; 0 other:  
 Query Match 1.0%; Score 24; DB 24; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+04;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1077 CTTTACTGCCACAAATCAGATTAA 1100  
 |||||  
 DB 34 CTTTACTGCCACAAATCAGATTAA 11

RESULT 5  
 AAV21495/C  
 ID AAV21495 standard; DNA: 31 BP.  
 AC AAV21495:  
 XX  
 DT 11-AUG-1998 (first entry)  
 XX  
 DE Plasmid pRSV-CREB341 3' CREB 283 primer.  
 XX  
 KW Protein-protein interaction; disruption; binding protein; modulator;  
 KW inhibitor; therapeutic; prophylactic; primer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9813502-A2.  
 XX  
 PD 02-APR-1998.  
 XX  
 PF 26-SEP-1997; 97WO-US17276.  
 XX  
 PR 27-SEP-1996; 96US-0721730.  
 XX  
 PA (ICOS-); ICOS CORP.  
 XX  
 PI Goodman RH, Hoekstra ME;  
 XX  
 DR WPI: 1998-230711/20.  
 XX  
 PT Identifying agents which disrupt protein/protein interactions -  
 PT using host cells transformed with DNA to form transcriptional  
 PT activating promoter which indirectly affects production of  
 PT selectable marker  
 XX  
 PS Example 1: Page 27; 96pp: English.  
 XX  
 CC Oligonucleotides AAV21487 and AAV21503 are used in the construction of  
 CC plasmids which are used in a method to identify agents which disrupt  
 CC protein/protein interactions using a host cell which is transformed or  
 CC transfected with DNA. The host cells are useful to demonstrate in vivo  
 CC binding capacity of both known and suspected binding partner proteins  
 CC in a recombinant system. The described expression system permits  
 CC systematic analysis of the structure and function of a particular  
 CC binding protein, thus permitting identification and/or synthesis of  
 CC potential modulators of the physiological activity of the binding  
 CC proteins. The system can be used to identify and improve molecules which  
 CC are capable of inhibiting specific and general protein/protein  
 CC interactions. Inhibitors identified by the methods can then be examined  
 CC for utility in vivo as therapeutic and/or prophylactic medicaments for  
 CC conditions associated with various protein/protein interactions. The  
 CC system provides a sensitive assay and provides a positive signal when a  
 CC protein/protein interaction is disrupted.

XX SQ Sequence 31 BP; 3 A; 10 C; 10 G; 8 T; 0 other:  
 Query Match 1.0%; Score 23.4; DB 19; Length 31;  
 Best Local Similarity 96.0%; Pred. No. 8.9e+04;  
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 903 CAGCCTGCTGAAGACGACGACGAA 927  
 |||||

DB 31 CAGCCTGCTGAAGACGACGACGAA 7

RESULT 7  
 AAZ66915  
 ID AAZ66915 standard; DNA: 47 BP.  
 XX  
 AC AAZ66915:  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Human map-related biallelic marker SEQ ID NO:1262.  
 XX  
 KW Human genome; biallelic marker; high density disequilibrium map;  
 KW genomic map; haplotype; phenotype; polymorphic base; genotyping;  
 KW haplotyping; hybridisation; identification; characterisation;  
 KW diagnosis; single nucleotide polymorphism; SNP; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(24,C)  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 PN WO9954500-A2.  
 XX  
 PD 28-OCT-1999.  
 XX  
 PF 21-APR-1999; 99WO-IB00822.  
 XX  
 PR 21-APR-1999; 98US-0082614.  
 XX  
 PR 23-NOV-1998; 98US-0109732.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Cohen D, Blumenfeld M, Chumakov I;  
 XX  
 DR WPI: 2000-013267/01.  
 XX  
 PT Novel biallelic markers used to construct a high density disequilibrium  
 PT map of the human genome  
 XX  
 PS Claim 1: Page 499; 2745pp: English.  
 XX  
 CC AAZ65654 to AAZ69578 represent human biallelic markers from the present  
 CC invention, which contain a polymorphic base at position 24 of their  
 CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification  
 CC primers for the biallelic markers. The biallelic markers of the  
 CC invention have a variety of uses: they can be used for high density  
 CC mapping of the human genome, and in complex association studies and  
 CC haplotyping studies which are useful in determining the genetic basis  
 CC for disease states. Compositions and methods of the invention can also  
 CC be useful for the identification of the targets for the development of  
 CC pharmaceutical agents and diagnostic methods, as well as the  
 CC characterisation of the differential efficacious responses to and side  
 CC effects from pharmaceutical agents acting on a disease as well as other  
 CC treatment.  
 CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297  
 CC and 3367, are not actually given a sequence in the Sequence Listing  
 CC from the present invention.  
 XX  
 SQ Sequence 47 BP; 23 A; 1 C; 8 G; 15 T; 0 other:  
 Query Match 1.0%; Score 23.4; DB 21; Length 47;  
 Best Local Similarity 73.2%; Pred. No. 9.8e+04;  
 Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1364 AAGAAGTAAATTTGTTTACTTGTAAATTTGATGGAGAAA 1404  
 |||||  
 DB 1 AAGAAGTAAATTTGTTTACTTGTAAATTTGATGGAGAAA 41



XX Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
 KW Protein therapy; vaccine; probe; diagnostic assay; detection;  
 KW Quantitation; restorative therapy; polymorphic; ds.  
 XX Homo sapiens.  
 XX WO200140521-A2.  
 XX 07-JUN-2001.  
 XX 30-NOV-2000; 2000WO-US32758.  
 XX 30-NOV-1999; 99US-0168138.  
 XX 29-NOV-2000; 2000US-0726173.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 PI WPI; 2001-356160/37.  
 DR Polymorphic nucleic acid sequences, useful in genetic testing and  
 PT therapy -  
 PS Claim 1; Page 169; 2653pp; English.  
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide  
 CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
 CC AA53114 to AA53129 represent peptides related to human polymorphic  
 CC polynucleotide sequences, the sequences can be used in gene and protein  
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression of polymorphic polypeptides.  
 CC For example, (I) may be used to treat disorders by rectifying mutations  
 CC or deletions in a patient's genome that affect the activity of  
 CC polypeptides by expressing inactive proteins or to supplement the  
 CC patients own production of polypeptide. Additionally, (I) and its  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids  
 CC in samples, and therefore which patients may be in need of restorative  
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
 CC production of antibodies specific for polymorphic polypeptides. The  
 CC antibodies may also be used to down regulate expression and activity.  
 CC The antibodies may also be used as diagnostic agents for detecting the  
 CC presence of polymorphic polypeptides in samples.  
 XX Sequence 50 BP; 15 A; 3 C; 6 G; 26 T; 0 other;  
 SQ

Query Match 0.9%; Score 22.4; DB 22; Length 50;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+05;  
 Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 2310 ATTGTCATCTTACAAACATTTTGTCTCTCTGTAAGAGTAG 2357  
 ||| | ||||| |||| |||| | || | || |||| |  
 Db 1 ATTTTCTTAAACAAATTTTGTCTCTCTCTGTAAGAGTAG 48

RESULT 11  
 AAT07082/C  
 ID AAT07082 standard; DNA; 45 BP.  
 XX  
 AC AAT07082;  
 XX  
 DT 02-JUL-1996 (first entry)  
 XX  
 DE Primer ENDO.1, amplifies RNase P conjugate.  
 XX  
 KW Primer: RNase P conjugate; amplification; PCR; Endo.P1; Endo.P2;  
 KW sequence specific cleavage; gene therapy; pathogenic RNA;  
 KW viral transcript; RNA genome; disease-causing mRNA; oncogene; ss.  
 XX  
 OS Synthetic.

XX WO9532283-A1.  
 PN 30-NOV-1995.  
 XX  
 XX 23-MAY-1995; 95WO-US06519.  
 PF 23-MAY-1994; 94US-0247776.  
 PR (INDV ) UNIV INDIANA FOUND.  
 XX  
 XX Frank DH, Harris ME, Pace NR;  
 PI WPI; 1996-020578/02.  
 DR  
 XX RNase P RNA conjugates function as sequence-specific endonuclease(s)  
 PF - useful for gene therapy and study cf. e.g. viral activity in  
 PT vitro.  
 PS Example; Page 14; 49pp; English.  
 XX  
 CC The sequences given in AAT07078-84 are primers which were used in the  
 CC construction of the RNase P conjugates of the invention. These  
 CC primers amplify the Endo.P1 and Endo.P2 genes. The conjugate also  
 CC includes an oligonucleotide which includes a predetermined sequence  
 CC complementary to and available for hybridization with a nucleotide  
 CC sequence of the target sequence. The conjugates cause sequence  
 CC specific cleavage of the target oligonucleotide. The conjugates are  
 CC useful in gene therapy for the selective degradation of pathogenic  
 CC RNA's such as viral transcripts or RNA genomes, or of disease-  
 CC causing mRNA's such as products of oncogenes.  
 XX Sequence 45 BP; 13 A; 9 C; 9 G; 14 T; 0 other;  
 SQ

Query Match 0.9%; Score 22.2; DB 17; Length 45;  
 Best Local Similarity 77.1%; Pred. No. 1.7e+05;  
 Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1823 TRACCACTGACGACAAAGCATGTGTTTGAATT 1857  
 ||||| ||||| |||| | |||| | |||| |  
 Db 38 TRACCACTGACGACGACGATGTGTTTGAATT 4

RESULT 12  
 ABL40867  
 ID ABL40867 standard; DNA; 33 BP.  
 XX  
 AC ABL40867;  
 XX  
 DT 03-JUL-2002 (first entry)  
 XX  
 DE CREB cDNA amplifying primer #1.  
 XX  
 KW Nucleic acid detection; transcription factor; cytostatic; reporter gene;  
 KW neuroprotective; cancer diagnosis; Li-Fraumeni syndrome; memory loss;  
 KW CREB; PCR primer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200229043-A1.  
 XX  
 PD 11-APR-2002.  
 XX  
 XX 28-SEP-2001; 2001WO-JP08576.  
 PF 02-OCT-2000; 2000JP-0302100.  
 PR (HELI-) HELIX RES INST.  
 XX  
 XX Oda T, Muramatsu M;  
 PI NFI; 2002-340106/37.  
 XX  
 DR  
 XX

PT Identifying target gene of transcription factor for isolation in cells  
PT with vectors expressing fused protein of transcription factor with its  
PT activator and specific reporter gene, useful in cancer diagnosis and  
PT therapy -  
PS Example 1: Page 12: 32pp: Japanese.  
XX The invention relates to a method of detecting the target gene of a  
XX specific transcription factor. The method involves (a) supplying cells  
XX containing a vector expressing a fused protein of the transcription  
XX factor with an activator of this transcription factor and another vector  
XX containing a reporter gene; (b) separating cells; and (c) comparing the  
XX gene expression dose of the cells. The method is for identifying a target  
XX gene of transcription factor for isolation, which is used for application  
XX in cancer diagnosis and drug development for e.g. Li-Fraumeni syndrome  
XX and long-term memory loss. With expression of the reporter gene as  
XX indication, the identification and isolation can be efficiently carried  
XX out because its enlarged expression dose in a cell is induced by the  
XX transcription factor. Sequences ABL40867-874 represents primers used in  
XX the course of the invention.  
SQ Sequence 33 BP; 10 A; 6 C; 8 G; 9 T; 0 other;

Query Match 0.94; Score 22; DB 24; Length 33;  
Best Local Similarity 83.3%; Pred. No. 1.8e+05;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 110 TACTAATGACCATGGAATCTGGAGCCGA 139  
D6 3 TGAATCATGACCATGGAATCTGGAGCAGA 32

RESULT 13  
ABL01753/c

ID ABL01753 standard; DNA; 46 BP.

XX ABL01753;

D7 18-MAR-2002 (first entry)

XX Human MSH2 (hMSH2) intronic sequence SEQ ID NO:106.

XX Human; MLH1; MSH2; hMLH1; hMSH2; variant gene; diagnosis; HNPCC;  
XX hereditary non-polyposis colorectal cancer; ds.

XX Homo sapiens.

XX US2001044936-A1.

XX 22-NOV-2001.

XX 22-OCT-1999; 99US-0426548.

XX 22-OCT-1998; 98US-1051E0P.

XX (ROBB/) ROBBINS D.

XX (LING/) LIN-GOERKE J L.

XX (LING/) LING J C.

XX Robbins D, Lin-Goerke JL, Ling JC;

XX WPI; 2002-105577/14.

XX New variants of the human MLH1 and MSH2 genes for diagnosing or  
XX determining a predisposition for hereditary non-polyposis colorectal  
XX cancer -

XX Disclosure; Page 4; 38pp; English.

XX The present invention describes a variant human MLH1 or MSH2 gene.  
XX Also described are: (1) a method for diagnosing or predicting  
XX susceptibility to hereditary non-polyposis colorectal cancer (HNPCC),  
XX comprising screening a DNA sample for the variant MLH1 or MSH2 gene

CC where presence of the variant indicates presence of, or susceptibility  
CC to HNPCC; (2) a method of identifying mutants in splice donor or  
CC acceptor sites of a human MLH1 gene, comprising sequencing splice donor  
CC or acceptor sites of the gene with intronic primers for the human MLH1  
CC gene and analysing the sequence to identify any mutants; (3) a method of  
CC identifying mutants in splice donor or acceptor sites of a human MSH2  
CC gene, comprising sequencing splice donor or acceptor sites of the gene  
CC with intronic primers for the human MSH2 gene and analysing the sequence  
CC to identify any mutants; and (4) a transgenic model system for  
CC colorectal cancer comprising cells expressing the variant MLH1 or MSH2  
CC gene. The hMLH1 and hMSH2 variants are used to diagnose or determine a  
CC patient's susceptibility to hereditary non-polyposis colorectal cancer.  
CC ABL01648 to ABL01745 and ABL01746 to ABL01831 represent hMLH1 and hMSH2  
CC gene fragments from the present invention. ABL01832 to ABL01839  
CC represent mutagenic primers used in the exemplification of the present  
CC invention.

SQ Sequence 46 BP; 14 A; 5 C; 4 G; 23 T; 0 other;

Query Match 0.98; Score 22; DB 24; Length 46;

Best Local Similarity 73.7%; Pred. No. 1.9e+05;

Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1870 AAGTAATTAAGTACAAAGCATTTTATTAGTACTAC 1907

D6 43 AAAAAAAGGATACATTACATTTGAGATATTAC 6

RESULT 14

AAI71996/c

ID AAI71996 standard; DNA; 47 BP.

XX AAI71996;

DT 21-FEB-2002 (first entry)

XX VCP1579 primer HIVP100.

XX Immunisation; Human Immunodeficiency Virus; HIV; envelope glycoprotein;  
XX priming antigen; booster antigen; attenuated; viral vector; primer; PCR;  
XX long terminal repeat; LTR; HIV-1; polymerase chain reaction; amplify; ss.

XX Synthetic.

XX WO200182962-A2.

XX 08-NOV-2001.

XX 25-APR-2001; 2001WO-CA00577.

XX 27-APR-2000; 2000US-200011P.

XX (AVET ) AVENTIS PASTEUR LTD.

XX Rovinski B, Tartaglia J, Cao S, Persson R, Klein MH;

XX WPI; 2002-034490/04.

XX Immunizing against Human Immunodeficiency Virus (HIV) using primary and  
XX booster antigens -

XX Example 4; Page 16; 38pp; English.

XX The sequences given in AAI71986-97 are primers which were used in  
XX the production of the recombinant poxvirus, VCP1579. VCP1579 contains  
XX the HIV-1 gag and protease genes derived from the HIV-1 isolate IIB,  
XX the gp120 envelope sequences derived from the HIV-1 Bx08 isolate, and  
XX sequences encoding a polypeptide encompassing the known human  
XX cytotoxic T lymphocytes (CTL) epitopes from HIV-1 Nef and Pol. VCP1579  
XX may be used to immunise against Human Immunodeficiency Virus (HIV). The  
XX method of the invention for immunising against HIV infection uses priming  
XX (DNA encoding an envelope glycoprotein of a primary HIV-1 isolate) and  
XX booster antigens (non-infectious, non-replicating immunogenic HIV-1-like

Matches	Conservative	Mismatches	Indels	Gaps
28	0	10	0	0



Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 ATACAGATTGCCACAGCACCCTAG 853  
|||||  
DB 27 ATACAGATTGCCACAGCACCCTAG 1

RESULT 2  
US-09-225-928-686/c  
; Sequence 685, Application US/09225928  
; Patent No. 6352829  
; GENERAL INFORMATION:  
; APPLICANT: Chenchik, Alex  
; Bibilashvili, Robert  
; Jekhadze, George  
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
; EXPRESSION  
; NUMBER OF SEQUENCES: 1375  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 2200 Sand Hill Road, Suite 100  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: 05-Jan-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/859,998  
; FILING DATE: 21-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Field, Bret E.  
; REGISTRATION NUMBER: 37,620  
; REFERENCE/DOCKET NUMBER: 09096/002001  
; TELEPHONE: 415-854-0875  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 686:  
; LENGTH: 27 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; OTHER INFORMATION: oligonucleotide primer  
US-09-225-928-686  
Query Match 1.1% Score 27; DB 4; Length 27;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 ATACAGATTGCCACAGCACCCTAG 853  
|||||  
DB 27 ATACAGATTGCCACAGCACCCTAG 1

RESULT 3  
US-08-859-998-685  
; Sequence 685, Application US/08859998  
; Patent No. 5994076  
; GENERAL INFORMATION:  
; APPLICANT: Chenchik, Alex  
; Jekhadze, George  
; Bibilashvili, Robert  
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
; EXPRESSION  
; NUMBER OF SEQUENCES: 1375  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 2200 Sand Hill Road, Suite 100  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

;; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
;; EXPRESSION  
;; NUMBER OF SEQUENCES: 1375  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson, P.C.  
;; STREET: 2200 Sand Hill Road, Suite 100  
;; CITY: Menlo Park  
;; STATE: CA  
;; COUNTRY: US  
;; ZIP: 94025  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: Windows95  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 21-MAY-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Field, Bret E.  
;; REGISTRATION NUMBER: 37,620  
;; REFERENCE/DOCKET NUMBER: 09096/002001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-854-0875  
;; TELEFAX: 415-854-0875  
;; INFORMATION FOR SEQ ID NO: 685:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 26 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
;; FEATURE:  
;; OTHER INFORMATION: oligonucleotide primer  
US-08-859-998-685  
Query Match 1.1% Score 26; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 GCACCAGGAGTCCCAAGGATTGAAGA 532  
|||||  
DB 1 GCACCAGGAGTCCCAAGGATTGAAGA 26

RESULT 4  
US-09-225-928-685  
; Sequence 685, Application US/09225928  
; Patent No. 6352829  
; GENERAL INFORMATION:  
; APPLICANT: Chenchik, Alex  
; Jekhadze, George  
; Bibilashvili, Robert  
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
; EXPRESSION  
; NUMBER OF SEQUENCES: 1375  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 2200 Sand Hill Road, Suite 100  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/225,928  
 FILING DATE: 05-Jan-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/859,998  
 FILING DATE: 21-May-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Field, Bret E.  
 REGISTRATION NUMBER: 37,620  
 REFERENCE/DOCKET NUMBER: 09096/002001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-322-5070  
 TELEFAX: 415-854-0875  
 INFORMATION FOR SEQ ID NO: 685:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 26 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 FEATURE:  
 OTHER INFORMATION: oligonucleotide primer  
 SEQUENCE DESCRIPTION: SEQ ID NO: 685:  
 US-09-225-928-685

Query Match 1.1%; Score 26; DB 4; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 7e+02;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 GCACCAGGAGTCCCAAGGATTGAAGA 532  
 DB 1 GCACCAGGAGTCCCAAGGATTGAAGA 26

RESULT 5  
 US-08-881-094-33  
 Sequence 33, Application US/08881094A  
 Patent No. 6022739  
 GENERAL INFORMATION:  
 APPLICANT: Ryan, Clarence A.  
 APPLICANT: Pearce, Gregory L.  
 APPLICANT: McGurl, Barry F.  
 TITLE OF INVENTION: Systemin  
 FILE REFERENCE: 7555-000001CPB  
 CURRENT APPLICATION NUMBER: US/08/881,094A  
 CURRENT FILING DATE: 1997-07-09  
 EARLIER APPLICATION NUMBER: 08/308,887  
 EARLIER FILING DATE: 1994-09-19  
 EARLIER APPLICATION NUMBER: PCT/US93/02428  
 EARLIER FILING DATE: 1993-03-18  
 EARLIER APPLICATION NUMBER: 07/885,412  
 EARLIER FILING DATE: 1992-03-19  
 EARLIER APPLICATION NUMBER: 07/528,956  
 EARLIER FILING DATE: 1990-05-25  
 EARLIER APPLICATION NUMBER: PCT/US91/03685  
 EARLIER FILING DATE: 1991-05-24  
 NUMBER OF SEQ ID NOS: 43  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 33  
 LENGTH: 40  
 TYPE: DNA  
 ORGANISM: Lycopersicon esculentum  
 US-08-881-094-33

Query Match 0.9%; Score 21; DB 3; Length 40;  
 Best Local Similarity 73.0%; Pred. No. 1.5e+04;  
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 503 TGATGCACGAGGTGCGCAAGGATTGAAGAAGAGAG 539  
 DB 4 TCATACACAGAAATACCAAGATGGAACATGAGAG 40

RESULT 6  
 US-07-977-696C-60  
 Sequence 60, Application US/07977696C  
 Patent No. 572852  
 GENERAL INFORMATION:  
 APPLICANT: do Couto, Fernando J.R.  
 APPLICANT: Ceriani Dr., Roberto L.  
 APPLICANT: Peterson Dr., Jerry A.  
 APPLICANT: Padlan Dr., Eduardo A.  
 TITLE OF INVENTION: Analogue Peptides with Specificity  
 TITLE OF INVENTION: for Carcinomas and Kit and Diagnostic Vaccination  
 TITLE OF INVENTION: and Therapeutic Methods.  
 NUMBER OF SEQUENCES: 81  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI  
 STREET: 444 South Flower Street, Suite 2000  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: USA  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/977.696C  
 FILING DATE: 11-16-92  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Anzel Ph.D., Viviana  
 REGISTRATION NUMBER: 30,930  
 REFERENCE/DOCKET NUMBER: P66 38227  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (510) 748-6868  
 TELEFAX: (510) 748-6868  
 TELEX: n.a.  
 INFORMATION FOR SEQ ID NO: 60:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 50 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-07-977-696C-60

Query Match 0.9%; Score 21; DB 1; Length 50;  
 Best Local Similarity 73.0%; Pred. No. 1.6e+04;  
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

CY 953 CAGGGAAGCAGCTCGAGACTGTCTGAGAAAGAGAA 989  
 DB 14 CAGGGAAGCGCTTGAGTGGTGTCTGAGAAATAGAA 50

RESULT 7  
 US-08-129-93CB-60  
 Sequence 60, Application US/08129930B  
 Patent No. 5804167  
 GENERAL INFORMATION:  
 APPLICANT: do Couto Dr., Fernando J.R.  
 APPLICANT: Ceriani Dr., Roberto L.  
 APPLICANT: Peterson Dr., Jerry A.  
 APPLICANT: Padlan Dr., Eduardo A.  
 TITLE OF INVENTION: Analogue Peptides with Broad  
 TITLE OF INVENTION: Carcinoma Specificity, and Kit and  
 TITLE OF INVENTION: Diagnostic Vaccination and  
 TITLE OF INVENTION: Therapeutic Methods  
 NUMBER OF SEQUENCES: 96  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: V. AMZEL & ASSOC.  
 STREET: 2055 No. 5804187th Broadway, Suite 201  
 CITY: Walnut Creek



```

:
: GENERAL INFORMATION:
:
: APPLICANT: Laboratory of Molecular Biophotonics
: TITLE OF INVENTION: Nucleic Acid Detection in Cytoplasm
: FILE REFERENCE: BB699-02
: CURRENT APPLICATION NUMBER: US/05/476,256
: CURRENT FILING DATE: 1999-12-30
: NUMBER OF SEQ ID NOS: 29
:
: SEQ ID NO 29
: LENGTH: 40
: TYPE: DNA
:
:

```

ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: probe  
US-09-476-256-29

Query Match 0.8% Score 20.6; DB 4; Length 40;  
Best Local Similarity 74.3%; Pred. No. 1.9e+04;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1329 GAAGAGACTTCTGCTTTTCAACCCCGCCCTCTC 1363  
DB 2 GATGAGAAGTCTGCTTTGCCCCCCCCCCCCCCCC 36

RESULT 11  
US-09-349-644-13  
Sequence 13, Application US/09349644A  
Patent No. 6303340

GENERAL INFORMATION:  
APPLICANT: Pollitt, Stephen M.  
APPLICANT: Buckley, Douglas I.  
APPLICANT: Statish, Peter A.  
APPLICANT: Hartman, Tawmar E.  
TITLE OF INVENTION: METHOD FOR PRODUCING A PEPTIDE WITH A PI  
FILE REFERENCE: SCIOS.019A  
CURRENT APPLICATION NUMBER: US/09/349.644A  
CURRENT FILING DATE: 1999-07-08  
EARLIER FILING DATE: 1999-07-08  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 44  
TYPE: DNA  
ORGANISM: E. coli  
US-09-349-644-13

Query Match 0.8% Score 20.6; DB 4; Length 44;  
Best Local Similarity 74.3%; Pred. No. 2e+04;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2112 TATATTTCTTTTGAATGTTACATGAGAAA 2146  
DB 3 TTTTATTTTAAATGTTATGTTACATGAGAAA 37

RESULT 12  
US-09-980-071-21/c  
Sequence 21, Application US/09980071  
Patent No. 5914318

GENERAL INFORMATION:  
APPLICANT: Baum, James A.  
APPLICANT: Gilmer, Amy Jelen  
APPLICANT: Mettus, Anne-Marie Light  
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING  
TITLE OF INVENTION: LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS  
NUMBER OF SEQUENCES: 76  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/980.071  
FILING DATE: Concurrently Herewith

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/757,536  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MECO:206  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-980-071-21

Query Match 0.8% Score 20.6; DB 2; Length 50;  
Best Local Similarity 67.4%; Pred. No. 2.1e+04;  
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1332 GAGACTTCTGCTTTTCAACCCCGCCCTCTC 1374  
DB 50 GACATTATGCTGTCTCAACCCCGCCCTCTCTCC 8

RESULT 13  
US-08-757-536-21/c  
Sequence 21, Application US/08757536  
Patent No. 5942664

GENERAL INFORMATION:  
APPLICANT: Baum, James A.  
APPLICANT: Gilmer, Amy Jelen  
APPLICANT: Mettus, Anne-Marie Light  
TITLE OF INVENTION: Bacillus thuringiensis CryIC  
TITLE OF INVENTION: Compositions Toxic to Lepidopteran Insects and Methods for  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White and Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/757.536  
FILING DATE: CONCURRENTLY HEREWITH  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: MOBT:023  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-757-536-21

Query Match 0.8% Score 20.6; DB 2; Length 50;  
Best Local Similarity 67.4%; Pred. No. 2.1e+04;

; TITLE OF INVENTION: Max

; TITLE OF INVENTION: Max

Search completed: March 9, 2003, 21:46:21  
Job time : 106 secs



; CURRENT APPLICATION NUMBER: US/09/842,883  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/200,011  
; PRIOR FILING DATE: 2000-04-27  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 47  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus  
US-09-842-883-13

Query Match 0.9%; Score 22; DB 10; Length 47;  
Best Local Similarity 73.7%; Pred. No. 5.9e+04;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 968 AGAGTCTCTAGAGAGAGAGATATGCAATGTT 1005  
DB 47 AGAGTCTCTAGAGAGAGAGATATGCAATGTT 10

## RESULT 3

US-09-842-883-14  
; Sequence 14, Application US/09842883  
; Patent No. US20020051770A1  
; GENERAL INFORMATION:  
; APPLICANT: Rovinski, Benjamin  
; APPLICANT: Tartaglia, James  
; APPLICANT: Cao, Shi-Xian  
; APPLICANT: Persson, Roy  
; APPLICANT: Klein, Michel H.  
; TITLE OF INVENTION: IMMUNIZING AGAINST HIV INFECTION  
; FILE REFERENCE: 1038-1142 MIS  
; CURRENT APPLICATION NUMBER: US/09/842,883  
; CURRENT FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 60/200,011  
; PRIOR FILING DATE: 2000-04-27  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 48  
; TYPE: DNA  
; ORGANISM: Human immunodeficiency virus  
US-09-842-883-14

Query Match 0.9%; Score 22; DB 10; Length 48;  
Best Local Similarity 73.7%; Pred. No. 6e+04;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 968 AGAGTCTCTAGAGAGAGAGATATGCAATGTT 1005  
DB 1 AGAGTCTCTAGAGAGAGAGATATGCAATGTT 38

## RESULT 4

US-09-838-386-17  
; Sequence 17, Application US/09838386  
; Patent No. US20010055756A1  
; GENERAL INFORMATION:  
; APPLICANT: Pellerin, Charles  
; APPLICANT: Kukolj, George  
; TITLE OF INVENTION: Internal De No. US20010055756A1o Initiation Sites of the HCV NS5B  
; FILE REFERENCE: 1011.2180001  
; CURRENT APPLICATION NUMBER: US/09/838,386  
; CURRENT FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 60/198,793  
; PRIOR FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 17  
; LENGTH: 45  
; TYPE: DNA

; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: oligonucleotide  
US-09-838-386-17

Query Match 0.8%; Score 20.8; DB 10; Length 45;  
Best Local Similarity 70.0%; Pred. No. 1.1e+05;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1168 AATAACATTTTATTTCTAACAATTTCTTTTCTATG 1207  
DB 3 AATAGGCCATTTTCTTTTCTTTTCTTTTCTTTG 42

## RESULT 5

US-09-838-386-18/c  
; Sequence 18, Application US/09838386  
; Patent No. US20010055756A1  
; GENERAL INFORMATION:  
; APPLICANT: Pellerin, Charles  
; APPLICANT: Kukolj, George  
; TITLE OF INVENTION: Internal De No. US20010055756A1o Initiation Sites of the HCV  
; FILE REFERENCE: 1011.2180001  
; CURRENT APPLICATION NUMBER: US/09/838,386  
; CURRENT FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 60/198,793  
; PRIOR FILING DATE: 2000-04-21  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 18  
; LENGTH: 45  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: oligonucleotide  
US-09-838-386-18

Query Match 0.8%; Score 20.8; DB 10; Length 45;  
Best Local Similarity 70.0%; Pred. No. 1.1e+05;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1168 AATAACATTTTATTTCTAACAATTTCTTTTCTATG 1207  
DB 43 AATAGGCCATTTTCTTTTCTTTTCTTTTCTTTG 4

## RESULT 6

US-10-027-806-92/c  
; Sequence 92, Application US/10027806  
; Patent No. US20020160476A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCORP 002A  
; CURRENT APPLICATION NUMBER: US/10/027,806  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 92  
; LENGTH: 43  
; TYPE: DNA  
; ORGANISM: Cenarchaeum symbiosum  
; FEATURE:  
; NAME/KEY: TATA\_signal  
; LOCATION: (11)...(16)  
US-10-027-806-92

```
Query Match      0.88; Score 20.4; DB 9; Length 43;
Best Local Similarity 71.18; Pred. No. 1.3e+05;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 90 GCAGGAGCTTTACACCGGTAACTAAATGACCATCGA 127
DB 39 GCAGCAGGTGTACCCCGGTAAATTAATGAGCCCGGA 2

RESULT 7
US-10-034-623-92/c
; Sequence 92, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; FILE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; TITLE REFERENCE: DEORP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 43
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: (11)...(16)
US-10-034-623-92

Query Match      0.88; Score 20.4; DB 9; Length 43;
Best Local Similarity 71.18; Pred. No. 1.3e+05;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 90 GCAGGAGCTTTACACCGGTAACTAAATGACCATCGA 127
DB 39 GCAGCAGGTGTACCCCGGTAAATTAATGAGCCCGGA 2

RESULT 8
US-09-853-526-206
; Sequence 206, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilyva, Chumakov
; APPLICANT: Bougueleret, Lydie
; FILE OF INVENTION: PROSTATE CANCER GENE
; TITLE REFERENCE: CENSEP.18CPCP
; CURRENT APPLICATION NUMBER: US/09/853,526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.m
; SEQ ID NO 206
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1)...(47)
; OTHER INFORMATION: polymorphic fragment 4-50-323
; NAME/KEY: allele
; LOCATION: (24)...(24)
; OTHER INFORMATION: polymorphic base C
; NAME/KEY: primer_bind
; LOCATION: (1)...(23)
; OTHER INFORMATION: potential microsequencing oligo 4-50-323.mis1
; NAME/KEY: primer_bind
; LOCATION: (25)...(47)
; OTHER INFORMATION: complement potential microsequencing oligo 4-50-323.mis2
US-09-853-526-206

Query Match      0.88; Score 20.4; DB 10; Length 47;
Best Local Similarity 65.28; Pred. No. 1.3e+05;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1060 TAAAGCACCTTAAGGACCTTTACTGCCACAATCAGATTAAATTTGG 1105
DB 2 TTAACACATTGATGAATCTTTACTACTACAAAAGGTTGCGATTAG 47

RESULT 9
US-09-901-484A-206
; Sequence 206, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; FILE OF INVENTION: Prostate Cancer Gene
; TITLE REFERENCE: GEN-T11XG3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1)...(47)
; OTHER INFORMATION: polymorphic fragment 4-50-323
; NAME/KEY: allele
; LOCATION: (24)...(24)
; OTHER INFORMATION: polymorphic base C
; NAME/KEY: primer_bind
; LOCATION: (1)...(23)
; OTHER INFORMATION: potential microsequencing oligo 4-50-323.mis1
; NAME/KEY: primer_bind
; LOCATION: (25)...(47)
; OTHER INFORMATION: complement potential microsequencing oligo 4-50-323.mis2
US-09-901-484A-206

Query Match      0.88; Score 20.4; DB 10; Length 47;
Best Local Similarity 65.28; Pred. No. 1.3e+05;
Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1060 TAAAGCACCTTAAGGACCTTTACTGCCACAATCAGATTAAATTTGG 1105
```

Db    2   TTTAAACATTGATGAATCTTCTACTACAAGGGTTCGATTAG    47

RESULT 10  
US-10-079-623-71/c  
; Sequence 71, Application US/10079623  
; Patent No. US20020169302A1  
; GENERAL INFORMATION:  
; APPLICANT: Havukkala, Iikka J.  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Grigor, Murray R.  
; APPLICANT: Molenaar, Adrian J.  
; TITLE OF INVENTION: Compositions isolated from bovine  
; FILE REFERENCE: 11000.1044c3  
; CURRENT APPLICATION NUMBER: US/10/079,623  
; CURRENT FILING DATE: 2002-02-19  
; NUMBER OF SEQ ID NOS: 370  
; SOFTWARE: FastSeq for Windows version 4.0  
; SEQ ID NO 71  
; LENGTH: 44  
; TYPE: DNA  
; ORGANISM: Bovine  
US-10-079-623-71  
Query Match                  0.8%; Score 20.2; DB 9; Length 44;  
Best Local Similarity      68.3%; Pred. No. 1.4e+05;  
Matches    28; Conservative         0; Mismatches 13; Indels         0; Gaps         0;

QY    1780 CAGTCTCCTATGTGCTCAACATTTTGTAATACATACTAAAAGA    1820  
         |||||    |||||    |||||    |||||    |||||    |||||  
Db    44 CAATCTCGATATGCITTCACCTTTGGTAATATATTTTAAAGA    4

RESULT 11  
US-09-827-289-22  
; Sequence 22, Application US/09827289  
; Patent No. US20020009716A1  
; GENERAL INFORMATION:  
; APPLICANT: Abarzua, Patricia  
; TITLE OF INVENTION: Process for Allele Discrimination Using Primer  
; FILE REFERENCE: 469290-55  
; CURRENT APPLICATION NUMBER: US/09/827,289  
; CURRENT FILING DATE: 2001-04-C5  
; PRIOR APPLICATION NUMBER: U.S. 60/194843  
; PRIOR FILING DATE: 2000-04-05  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 46  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: P1 primer for  
; OTHER INFORMATION: use in allele discrimination  
US-09-827-289-22

Query Match                  0.8%; Score 20; DB 10; Length 46;  
Best Local Similarity      72.2%; Pred. No. 1.6e+05;  
Matches    26; Conservative         0; Mismatches 10; Indels         0; Gaps         0;

QY    1176 TTTTATTTCTTAAACATCTCTTTTTTTCTATCGCGA    1211  
         |||||    |||||    |||||    |||||    |||||    |||||  
Db    11 TTTTATTTCTTAAACATCTCTTTTTTTCTATCTCTCA    46

RESULT 12  
US-09-864-795-3523  
; Sequence 3523, Application US/09864795  
; Patent No. US20020177568A1  
; GENERAL INFORMATION:

Patent No. US20020150891A1  
GENERAL INFORMATION:  
APPLICANT: Hood, Leroy E.  
APPLICANT: Roven, Ben F.  
APPLICANT: Koop, Ben F.  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH  
NUMBER OF SEQUENCES: 1279  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/263,959  
FILING DATE: 05-MAR-1999  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: McHesters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 920010.426C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 486:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-263-959-486  
Query Match: 0.8%; Score 19.8; DB 10; Length 45;  
Best Local Similarity 77.4%; Pred. No. 1.7e+05;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
OY 1176 TTTTATTTCTTAACATTTCTTTTCTAT 1206  
DB 2 TTTCTTTTCTTTCTTTCTTTTCTTTT 32  
Search completed: March 9, 2003, 21:49:51  
Job time : 199 secs

LENGTH: 41 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-729-402-28  
Query Match: 0.8%; Score 19.8; DB 10; Length 41;  
Best Local Similarity 69.2%; Pred. No. 1.5e+05;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
OY 707 TGCAGCAGCCACTCAGCCGGTACTACCATTTCTACAGTA 745  
DB 39 TGCAGCAGCCACTTAATCTGTATCATTAATTTAGATA 1  
RESULT 14  
US-09-921-398-18  
Sequence 18, Application US/09921398  
Patent No. US20020055169A1  
GENERAL INFORMATION:  
APPLICANT: Tekamp-Olson, Patricia  
TITLE OF INVENTION: METHOD FOR EXPRESSION OF HETEROLOGOUS  
PROTEINS IN YEAST  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell Seltzer IP Group of Alston & Bird, LLP  
STREET: 3605 Glenwood Ave. Suite 310  
CITY: Raleigh  
STATE: NC  
COUNTRY: US  
ZIP: 27622  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 02-Aug-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5784-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919 420 2202  
TELEFAX: 919 881 3175  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-921-398-18  
Query Match: 0.8%; Score 19.8; DB 10; Length 45;  
Best Local Similarity 69.2%; Pred. No. 1.7e+05;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
OY 798 GTTCAAGTTCCTGTCGAGCTACAAATACAGATT 836  
DB 2 GTTCAAGTTCCTGTCGAGCTACAAATTCCTAGCTT 40  
RESULT 15  
US-09-263-959-486  
Sequence 486, Application US/09263959



OK nucleic - nucleic search, using sw model

Run on: March 9, 2003, 18:54:42 ; Search time 3109 Seconds  
(without alignments)  
12783.443 Million cell updates/sec

Title: US-09-973-827-3

Perfect score: 2454

Sequence: 1 cggcgaggaggttagtttg.....totgtatttgcataatcat 2454

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 102860

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 04

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pin:\*
- 21: em\_gss\_vit:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_nam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	1.8	50	9 AUI07544	AUI07544 AUI07544
2	22.8	0.9	50	17 AZ663650	AZ663650 LM0543C23
C 3	22.6	0.9	46	9 AUI014208	AUI014208 AUI014208
C 4	22.4	0.9	48	17 AZ361863	AZ361863 LM0106B17
C 5	22.2	0.9	46	9 AUI014238	AUI014238 AUI014238
C 6	22	0.9	50	9 AA545213	AA545213 vj93e11.r

C 7	22	0.9	50	13 BI425819	BI425819 sah72q08.
C 8	21.8	0.9	44	17 AL757113	AL757115 Arabidops
C 9	21.8	0.9	46	17 AZ820550	AZ820550 2M0092M18
C 10	21.6	0.9	48	13 BJ083775	BJ083775 BJ083775
C 11	21.4	0.9	46	10 AV851679	AV851679 AV851679
C 12	21.2	0.9	43	17 AZ638328	AZ638328 LM0498110
C 13	21.2	0.9	45	9 AUI014228	AUI014228 LM0498110
C 14	21	0.9	46	10 AV947975	AV947975 AV947975
C 15	21	0.9	50	10 AW333744	AW333744 S25E9 AGS
C 16	20.8	0.8	50	10 AW333744	AW333744 S25E9 AGS
C 17	20.8	0.8	46	17 AL760388	AL760388 Arabidops
C 18	20.6	0.8	49	17 BH861575	BH861575 SALK_0875
C 19	20.4	0.8	43	17 AZ435166	AZ435166 LM0229J07
C 20	20.4	0.8	45	2 HSM002460	AIO38111 Homo sapi
C 21	20.4	0.8	46	13 BJ015738	BJ015738 BJ015738
C 22	20.4	0.8	46	13 BJ075979	BJ075979 BJ075979
C 23	20.4	0.8	50	9 AJ499559	AJ499559 AJ499559
C 24	20.4	0.8	50	13 BJ066224	BJ066224 BJ066224
C 25	20.2	0.8	45	9 AU269673	AU269673 AU269673
C 26	20.2	0.8	45	17 AL761682	AL761682 Arabidops
C 27	20.2	0.8	48	17 HSC06506	XB9550 H. sapiens D
C 28	20.2	0.8	50	17 BH791446	BH791446 SALK_0599
C 29	20	0.8	43	9 AA509550	AA509550 vH50D03.r
C 30	20	0.8	43	17 AZ620145	AZ620145 LM0452D15
C 31	20	0.8	44	14 H98332	H98332 Y79033.s1
C 32	20	0.8	45	17 AZ427755	AZ427755 LM0209N24
C 33	20	0.8	46	17 AZ766403	AZ766403 LM0563023
C 34	20	0.8	48	17 AZ592310	AZ592310 LM0403D12
C 35	20	0.8	49	17 TA98D04Q	AL459661 T. brucei
C 36	20	0.8	50	9 AUI03398	AUI03398 AUI03398
C 37	19.8	0.8	32	17 AL761349	AL761349 Arabidops
C 38	19.8	0.8	40	17 AQ073791	AQ073791 EP(3)3174
C 39	19.8	0.8	40	17 TA105A11P	AL464696 T. brucei
C 40	19.8	0.8	46	9 AA946759	AA946759 q550c05.s
C 41	19.8	0.8	46	10 AW334262	AW334262 S32E7 AGS
C 42	19.8	0.8	46	13 BI256739	BI256739 602974296
C 43	19.8	0.8	46	17 AL770038	AL770038 Arabidops
C 44	19.8	0.8	47	17 AZ843462	AZ843462 2M0143P18
C 45	19.8	0.8	48	2 HSM002436	AIO38087 Homo sapi

ALIGNMENTS

RESULT 1  
AUI07544  
LOCUS AUI07544 Sugano Homo sapiens CDNA library Homo sapiens cDNA clone  
DEFINITION NBLAN589NF, mRNA sequence.  
ACCESSION AUI07544  
VERSION AUI07544.1 GI:13557065  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata  
H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki  
Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoka, Tokyo 108-8639, Japan  
Email: yusuk@iims.u-tokyo.ac.jp  
Suzuki,Y., Yoshimoto-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano  
S. Construction and characterization of a full length-enriched and  
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).  
Location/Qualifiers

AUI07544 50 bp mRNA linear EST 30-AUG-2001  
AUI07544 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
NBLAN589NF, mRNA sequence.

AUI07544.1 GI:13557065  
EST

human.  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata  
H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki  
Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoka, Tokyo 108-8639, Japan  
Email: yusuk@iims.u-tokyo.ac.jp  
Suzuki,Y., Yoshimoto-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano  
S. Construction and characterization of a full length-enriched and  
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).  
Location/Qualifiers

source 1. .50  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="WELAN359NF"  
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"  
BASE COUNT 8 a 6 c 25 g 11 t  
ORIGIN

Query Match 1.6%; Score 44; DB 9; Length 50;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCGCGAGGTGTAGTTTACGCGGTGTGTACGTGGGGAGA 44  
|||||  
Db 7 CCGCGCGAGGTGTAGTTTACGCGGTGTGTACGTGGGGAGA 50  
|||||

RESULT 2  
A2663650  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

A2663650  
1M0543G23F Mouse 10kb plasmid UUGCLM library Mus musculus genomic  
clone UUGCLM0543G23 F, DNA sequence.  
ACCESSION A2663650  
VERSION GSS  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 50);  
Dunn, P., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rallly,  
M., Rose, M., Ross, R., Stokes, R., Tingey, A., von Niederhausen, A.,  
and Wright, B., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
Plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0543 row: G column: 23  
Seq primer: CGTGTAAACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 50.  
Location/Qualifiers  
1. .50  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGCLM0543G23"  
/clone\_lib="Mouse 10kb plasmid UUGCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, P-"  
/note="Vector: PMD42uv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (914732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."  
BASE COUNT 15 a 17 c 5 g 13 t  
ORIGIN

Query Match 0.9%; Score 22.8; DB 17; Length 50;  
Best Local Similarity 71.4%; Pred. No. 1.6e+06;  
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 795 GTTGTTCAGCTCCCTGCGAGCGTACAAACATACAGATT 836  
|||||  
Db 1 GTCTTGCACACTCTCTGGACCTACACATACAGATT 42  
|||||

RESULT 3  
AU014208/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

AU014208  
AU014208 Schizosaccharomyces pombe late log phase cDNA  
Schizosaccharomyces pombe cDNA clone spc09381, mRNA sequence.  
ACCESSION AU014208  
VERSION EST  
KEYWORDS fission yeast.  
SOURCE Schizosaccharomyces pombe  
ORGANISM Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
Schizosaccharomycetales; Schizosaccharomycetaceae;  
Schizosaccharomyces.  
1 (bases 1 to 46)  
Morimyo, M. and Mita, K.  
Identification of expressed sequence tags of Schizosaccharomyces  
pombe  
Unpublished (1998)  
Contact: Mitsuki Morimyo  
Genome Research Group  
National Institute of Radiological Sciences  
9-1, Anagawa-4-chome, Inage-ku, Chiba 263-8555, Japan  
Email: morimyo@nirs.go.jp.  
Location/Qualifiers  
1. .46  
/organism="Schizosaccharomyces pombe"  
/strain="972"  
/db\_xref="taxon:4896"  
/clone="spc09381"  
/clone\_lib="Schizosaccharomyces pombe late log phase cDNA"  
/sex="h minus"  
/note="Vector: M13mp19; The cDNA library of  
Schizosaccharomyces pombe was prepared by cloning cDNA  
into the SmaI site of M13mp19 DNA and the direction of DNA  
sequences was not always from 5' to 3'. The cDNA data of  
Schizosaccharomyces pombe are available for searching on  
the World Wide Web. (URL, http://www.nirs.go.jp)"  
BASE COUNT 22 a 5 c 3 g 15 t 1 others  
ORIGIN

Query Match 0.9%; Score 22.6; DB 9; Length 46;  
Best Local Similarity 67.4%; Pred. No. 1.8e+06;  
Matches 31; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2122 TTATGAATGTACATGTGAGAAAATATCTGATTTTAAATATTTTC 2167  
|||||  
Db 46 TTTTTCACGTAAAGGTAGTATTATATTTGATTTTATATAAATTC 1  
|||||

RESULT 4  
A2361863  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source

A2361863  
1M0106B17R Mouse 10kb plasmid UUGCLM library Mus musculus genomic  
clone UUGCLM0106B17 R, DNA sequence.  
ACCESSION A2361863  
VERSION GSS  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 50);  
Dunn, P., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rallly,  
M., Rose, M., Ross, R., Stokes, R., Tingey, A., von Niederhausen, A.,  
and Wright, B., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
Plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0543 row: G column: 23  
Seq primer: CGTGTAAACGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 50.  
Location/Qualifiers  
1. .50  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGCLM0543G23"  
/clone\_lib="Mouse 10kb plasmid UUGCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, P-"  
/note="Vector: PMD42uv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative

ACCESSION AU014238  
 VERSION AU014238.1 GI:3369029  
 KEYWORDS EST.  
 ORGANISM fission yeast.  
 SOURCE Schizosaccharomyces pombe  
 ORGANISM Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 Schizosaccharomycetales; Schizosaccharomycetaceae;  
 Schizosaccharomyces.  
 1 (bases 1 to 46)  
 Morimyo,M. and Mita,K.  
 Identification of expressed sequence tags of Schizosaccharomyces  
 pombe  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Mitsuaki Morimyo  
 Genome Research Group  
 National Institute Of Radiological Sciences  
 9-1, Anagawa 4 chome, Inage-ku, Chiba, Chiba 263-8555, Japan  
 Email: morimyo@nirs.go.jp.  
 Location/Qualifiers  
 1. .46  
 /organism="Schizosaccharomyces pombe"  
 /strain="972"  
 /db\_xref="taxon:4896"  
 /clicne="spc09421"  
 /clicne.lib="Schizosaccharomyces pombe late log phase cDNA"  
 /sex="h minus"  
 /note="Vector: M13mp19; The cDNA library of  
 Schizosaccharomyces pombe was prepared by cloning cDNA  
 into the SmaI site of M13mp19 DNA and the direction of  
 sequences was not always from 5' to 3'. The cDNA data of  
 Schizosaccharomyces pombe are available for searching on  
 the World Wide Web. (URL, http://www.nirs.go.jp)"  
 22 a 3 c 3 g 13 t 3 others  
 BASE COUNT 22 a 3 c 3 g 13 t 3 others  
 ORIGIN  
 Query Match 0.98; Score 22.2; DB 9; Length 46;  
 Best Local Similarity 65.2; Pred. No. 2.2e+06;  
 Matches 30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 QY 2122 TTTTGAATGTTACATGTAGAAAATCTAGTTTAAATATATTC 2167  
 Db 46 TTTTTCACGTGAAGTAGTATTATATGATTTATNAATATTC 1  
 RESULT 6  
 AA545213/c  
 LOCUS  
 DEFINITION IMAGE:944584 3', mRNA sequence.  
 ACCESSION AA545213  
 VERSION AA545213.1 GI:2306287  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 50)  
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouche,T.,  
 Getzel,S., Kuback,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
 Schellenberg,K., Steptoe,K., Tan,F., Underwood,K., Moore,B.,  
 Thelsing,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and  
 Watson,R.  
 The WashU-HMNI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMNI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

```

/clone_lib="Gm-C1049"
/tissue_type="whole seedlings of greenhouse grown plants"
/dev stage="3 week old"

```



```

RESULT 12
A2638328
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

A2638328          43 bp   DNA          linear      GSS 13-DEC-2000
1M0498110         Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0498110 F. DNA sequence.
A2638328
A2638328.1       GI:11760518
GSS.
house mouse.
house musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 43)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hanil,C.,
Islam,B., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,F., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffold with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Em. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCT, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0498 row: I column: 10
,Seq primer: CCGTGTAAACACGCCGCGAT

```

/note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, http://www.nirs.go.jp)"

BASE COUNT 21 a 5 c 14 t 2 others

ORIGIN Query Match 0.98; Score 21.2; DB 9; Length 45; Best Local Similarity 67.48; Pred. No. 3.6e+06; Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 2122 TTTTGAATGTTACATGTCAGAAATACCTGATTTTAAATATT 2164

Db 45 TTTTTCACCTGTAAGGTAGTATATATTGATTTTATANAAT 3

RESULT 14

AV947975

LOCUS AV947975

DEFINITION AV947975 Nori Satoh unpublished cDNA library, young adult Ciona

ACCESSION AV947975

VERSION AV947975.1 GI:19425734

KEYWORDS EST.

SOURCE Ciona intestinalis.

ORGANISM Ciona intestinalis

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

AUTHORS Phlebobranchia; Clonidae; Ciona.

TITLE 1 (bases 1 to 46)

JOURNAL Satoh, N., Satoh, Y., Kohara, Y. and Shino, I. T.

COMMENT Expressed genes in Ciona intestinalis

Unpublished (2000)

Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

1. .46

/organism="Ciona intestinalis"

/db\_xref="taxon:7719"

/clone="ciad05102"

/clone\_lib="Nori Satoh unpublished cDNA library, young adult"

/tissue\_type="whole animal"

/dev\_stage="young adult"

/note="Vector: pBluescript SK"

18 a 0 c 4 g 22 t 2 others

BASE COUNT

ORIGIN

Query Match 0.98; Score 21; DB 10; Length 46;

Best Local Similarity 69.28; Pred. No. 3.9e+06;

Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1176 TTTTATTTCTAAACATTTCTTTTCTATGCGCAAA 1214

Db 2 TTTTATTTCTAAACATTTCTTTTCTATGCGCAAA 40

RESULT 15

AW333744

LOCUS AW333744

DEFINITION S25E9 AGS-1 Pneumocystis carinii f. sp. carinii linear EST 31-JAN-2000

ACCESSION AW333744

VERSION AW333744.1 GI:6830101

KEYWORDS EST.

SOURCE Pneumocystis carinii f. sp. carinii.

ORGANISM Pneumocystis carinii f. sp. carinii

REFERENCE 1 (bases 1 to 50)

AUTHORS Smullan, A. G., Arnold, J., Weise, M., Wunderlich, J., Staben, C., Edman, J. C., Kovacs, J. and Cushion, M.

TITLE Expressed sequence tags from Pneumocystis carinii

JOURNAL Unpublished (2000)

COMMENT Contact: Staben, C.

School of Biological Sciences

University of Kentucky

101 Morgan Building, University of Kentucky, Lexington, KY

40506-0225, USA

Tel: 606 257 2161

Fax: 606 257 1717

Email: staben@pop.uky.edu.

Location/Qualifiers

1. .50

/organism="Pneumocystis carinii f. sp. carinii"

/db\_xref="taxon:38081"

/clone\_lib="AGS-1"

/lab\_host="E. coli"

/note="Vector: Lambda ZAP II; Site 1: EcoRI; Site 2: XhoI;

P. carinii organisms (3x10<sup>8</sup>) from a single rat (99-1-6,

sacrificed on 3/17/99) at Cincinnati VA facilities.

Trizol extracted RNA. Oligo dT priming, standard

conditions described by vendor, Stratagene. Further

details see www.uky.edu/project/Pneumocystis/

BASE COUNT 21 a 1 c 1 g 27 t

ORIGIN

Query Match 0.98; Score 21; DB 10; Length 50;

Best Local Similarity 66.7%; Pred. No. 3.9e+06;

Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 2330 TTTTGTCTCTCTGTAAGAGAGTAGTATTAGTCTGCTTAA 2374

Db 4 TTTTGTCTCTCTGTAAGAGAGTAGTATTAGTCTGCTTAA 48

Search completed: March 9, 2003, 21:44:26

Job time : 3114 secs